

Piatek MJ, Henderson V, Zynad HS, Werner A.

[Natural antisense transcription from a comparative perspective.](#)

Genomics 2016, 108(2), 56-63.

Copyright:

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

DOI link to article:

<http://dx.doi.org/10.1016/j.ygeno.2016.05.004>

Date deposited:

20/10/2016



This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)



Natural antisense transcription from a comparative perspective[☆]



Monica J. Piatek, Victoria Henderson, Hany S. Zynad, Andreas Werner^{*}

RNA Interest Group, Institute for Cell and Molecular Biosciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, United Kingdom

ARTICLE INFO

Article history:

Received 16 February 2016

Received in revised form 8 May 2016

Accepted 25 May 2016

Available online 27 May 2016

Keywords:

Natural antisense transcript

Evolution

Na-phosphate transport

Human

Mouse

ABSTRACT

Natural antisense transcripts (NATs) can interfere with the expression of complementary sense transcripts with exquisite specificity. We have previously cloned NATs of *Slc34a* loci (encoding Na-phosphate transporters) from fish and mouse. Here we report the cloning of a human *SLC34A1*-related NAT that represents an alternatively spliced *PFN3* transcript (Profilin3). The transcript is predominantly expressed in testis. Phylogenetic comparison suggests two distinct mechanisms producing *Slc34a*-related NATs: Alternative splicing of a transcript from a protein coding downstream gene (*Pfn3*, human/mouse) and transcription from the bi-directional promoter (*Rbpja*, zebrafish). Expression analysis suggested independent regulation of the complementary *Slc34a* mRNAs. Analysis of randomly selected bi-directionally transcribed human/mouse loci revealed limited phylogenetic conservation and independent regulation of NATs. They were reduced on X chromosomes and clustered in regions that escape inactivation. Locus structure and expression pattern suggest a NATs-associated regulatory mechanisms in testis unrelated to the physiological role of the sense transcript encoded protein.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Natural antisense transcripts (NATs) are long regulatory RNAs that are transcribed in the opposite direction of protein-coding transcripts and display extended sequence complementarity [1]. A regulatory cascade can unfold in *cis* at the bi-directionally transcribed locus or in *trans* via base-pairing of the processed, complementary sense/antisense transcripts. Various regulatory mechanisms involving natural antisense transcripts have been identified, remarkably in genes that, apart from being bi-directionally transcribed, have not much in common. The genome-wide scale of antisense transcription became first apparent in large scale cDNA sequencing projects and bioinformatics approaches [2–4]. It is now recognized that a significant proportion of mammalian genes are transcribed in both directions; the estimates for a particular species depend crucially on sequencing depth and annotation quality of the relevant genomes and transcriptomes [5]. For both human and mouse there is evidence that 50% or more of all genomic loci are bi-directionally transcribed [2].

Bi-directional transcription is found in the genomes of most organisms, however, species specific regulatory strategies have evolved. In prokaryotes antisense transcripts act post-transcriptionally and interfere with secondary structures of complementary RNAs. These interactions ‘melt’ either inhibitory or stimulatory complexes leading to enhanced or repressed gene expression. An additional well documented

regulatory mechanism involves so-called ‘kissing complexes’ with short RNA-RNA interactions [6]. Many bacterial NATs have been characterized in detail and are extensively reviewed [7].

In eukaryotes sense-antisense interactions involve extended RNA-RNA complementarity. The potential formation of double-stranded RNA (dsRNA) resembles viral structures and poses a deadly threat to the cell [8]. Therefore, co-expression of sense-and antisense transcripts may trigger a defensive response rather than a regulatory cascade. Specific strategies have evolved to mitigate this dilemma: Plants, for example have entered a veritable arms race with viruses involving components of the RNA interference pathway [9]. In addition, endogenous siRNAs originating from sense-antisense duplexes are abundant and readily detectable in plant transcriptome sequencing projects. Vertebrates, on the other hand, fight viral structures using a dsRNA-specific kinase that elicits an immune response [10]. Moreover, dsRNA is an essential intermediate in genome defence strategies against transposable elements. Both pathways are well described [11]. In contrast, the biological relevance of dsRNA formed between sense-and antisense transcripts in vertebrates is still poorly understood and only few examples have been studied in detail [12–14].

NATs in animals show development and tissue specific regulation with significant enrichment in testis and brain [15,16]. Several lines of evidence confirm a co-expression of sense and antisense transcripts, though the NAT is generally expressed at a much lower level than the protein coding sense RNA. Whether the complementary transcripts form dsRNA is controversial: The evidence is either indirect or blurred through experimental complications (see Discussion). Nonetheless, it is plausible that sense-antisense dsRNA structures are being formed in animal cells, likely during small developmental windows, in specific

[☆] Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries (Submission 6 May 2016, No KX215110).

^{*} Corresponding author.

E-mail address: Andreas.werner@ncl.ac.uk (A. Werner).

cell types or in particular cellular compartments. An interesting genome-wide feature represents the significant under-representation of NATs on the mouse and human X chromosome, which may be related to the generally mono-allelic expression of X chromosomal genes [3,4].

The *Slc34a* gene family encodes epithelial Na-phosphate transporters that are essential in maintaining homeostasis of inorganic phosphate in most animals [17]. Their expression is tightly controlled by hormones and metabolic factors. Many *Slc34a* genes are transcribed in both orientations resulting in natural antisense transcripts complementary to a few exons of the protein-coding sense transcript. The *Slc34a*-related NATs characterized so far are fully processed, however, different start sites and splicing strategies have evolved; the common denominator appears to be substantial exonic complementarity [18–20]. Here, we report the cloning of a human *SLC34A1* antisense transcript. We use the publicly available databases to characterize the transcript and trace its evolutionary development. Moreover, we compare these findings to other well characterized sense/antisense pairs to gain insights into common patterns of gene regulation by NATs.

2. Material and methods

2.1. Molecular biology

Human testis RNA was purchased from Clontech/TaKaRa and used for 3'RACE (rapid amplification of cDNA ends) and RT-PCR without further purification. 3'RACE was performed using an oligo dT adaptor primer and 3 nested primers that all localize to the 5' end of the *PFN3* gene. The resulting fragment (780 bp) was cloned using the pGEM®-T vector system (Promega) and positive clones were identified by colony PCR. The sequence of the insert was determined by Sanger sequencing (Genevision, Newcastle, UK).

HEK293 cells were a kind gift from Dr. Nicolas Watkins (Newcastle University) and grown according to standard protocols. Total RNA was extracted using Trizol (Invitrogen) and treated with DNaseI (Thermo Scientific). Approximately 0.5 µg of total RNA was used for reverse transcription using the Omniscript kit (Qiagen). RNA and random hexamers (2.5 µM) were denatured for 3 min at 70 °C and cooled to 37 °C. Polymerase, dNTPs, buffer and RNase inhibitor were added as detailed by the supplier. The final reaction was denatured for 2 min at 95 °C and stored at –20 °C. Gene specific PrimeTime® qPCR primers and probes (Integrated DNA Technologies) were used to amplify a specific product from 0.5 µl of the RT reaction. 1 × Lightcycler 480 Probes Master mix (Roche) and with the cycling protocol: initial denaturation (95 °C for 10 min), 45 cycles of 95 °C for 5 s, 55 °C for 20 s and 72 °C for 1 s when fluorescence was determined. The sequence of all primers and probes are given as supplementary material (Supplementary Table 1).

2.2. Bioinformatics

For the phylogenetic analysis of human and mouse bi-directionally transcribed genes the Ensembl genome browser was used with the Ensembl-Havana human GENCODE gene set (release 24) and mouse (release 83). The loci were selected by manually picking random areas (10 on each autosome) and selecting the closest pair of complementary transcripts. Protein coding and processed transcripts were included, pseudogenes and RNA genes excluded. Transcripts were deemed to be complementary when 20 nucleotides or more were in common, with the assumption that RNA duplexes of >20 bp are likely to be stable under physiological conditions and biologically active. The loci were visually categorized with respect to overlap pattern (5', 3' or mid, Fig. 3). The EBI gene expression atlas was used to determine level and expression pattern of the individual transcripts (<http://www.ebi.ac.uk/gxa/home>). Peak expression as given by the expression atlas was used for the comparative analysis. (This choice was prior validated for 16 human and 8 mouse genes by normalizing individual data sets to the

highest value (100%); the averaged expression levels were then used to calculate a scaling factor for comparison between sense and antisense expression. The expression ranking and pattern was identical for both approaches.) The expression pattern was manually assessed and grouped into 'brain', 'testis' and 'others'. The two highest expressing tissues were selected (only one if the second highest expression was <60% of the highest value) and colour-coded in an excel table (Fig. 4, Supplementary Table 2). To estimate expression patterns, a value of two was given to transcripts expressed in a single tissue and a value of one each if the transcript was found at high levels (>60% of peak expression) in two or more tissues. This approach was validated using the scaled expression levels of 24 transcripts. All these data were compiled in spread sheets and excel functions were used to compare and sort the data according to the parameters indicated in the Results section. The names of genes that escape X chromosome inactivation were downloaded from Zhang et al. [21] and Berletch et al. [22].

3. Results

3.1. Cloning of the human *SLC34A1* antisense transcript

To demonstrate the expression of an *SLC34A1*-related antisense transcript we used total RNA from human testis; we chose this tissue because NATs generally show highest expression levels in mammalian testis, including the mouse homologue of the candidate gene [18]. The fact that the mouse NAT represents an alternative splice product of a *Pfn3* transcript further informed our search strategy (Fig. 1A). All cloned *SLC34A*-related NATs (from flounder, zebrafish and mouse [18–20]) are relatively short and alternatively spliced. We therefore applied 3' rapid amplification of cDNA ends (3'RACE) in combination with *PFN3* specific primers to amplify fragments of the human *SLC34A1* NAT. Three rounds of PCR with one specific primer each were necessary to amplify a specific fragment. Sanger sequencing revealed a transcript of 780 nucleotides containing 3 exons. The first one represents the start of *PFN3* and the second and third one are complementary to exon 13 of *SLC34A1* (Fig. 1A). A short open reading frame of 63 amino acids contains 47 N-terminal residues of Profilin 3 and 16 amino acids encoded by the noncoding 3' end of the *SLC34A1* transcript.

RT-qPCR, RT-PCR and publicly available expression data (<http://www.proteinatlas.org>) were used to investigate the expression pattern of both *SLC34A1*-related sense and antisense transcripts (Figs. 1B–D). Both transcripts showed a very restricted expression pattern with the expected dominant expression of the sense transcript in kidney, traces were found in testis but none in other tissues. The antisense transcript was found in testis at a low and in kidney at a hardly detectable level. Interestingly, the antisense transcript was also expressed in HEK293 cells without detectable levels of the protein coding sense RNA. We were unable to localize the *SLC34A1* antisense transcript on human kidney sections by *in situ* hybridization (not shown).

3.2. Phylogeny of *Slc34a*-related antisense transcripts

Antisense transcripts have been identified and cloned from different *Slc34a* isoforms. We used the comparative genomics tools from Ensembl to build a phylogenetic tree of *Slc34a* (Fig. 2A) and to visualize the region downstream where potential antisense transcripts may originate (Fig. 2B). The *Slc34a* gene family initially comprised two members in the animal kingdom, denoted *Slc34a1* and *Slc34a2*. Through contractions and gene duplications lineage specific genotypes have evolved. For example, in mammals a third isoforms is expressed in kidney, *Slc34a3*, and the protein shows distinct functional properties. On the other hand, a documented gene duplication in zebrafish has resulted in two *Slc34a1* and two *Slc34a2* isoforms. The gene arrangements downstream of the *Slc34a* loci predict that the *Slc34a*-linked antisense transcripts cloned and analysed to date (*Slc34a1* in mammals and *Slc34a2*

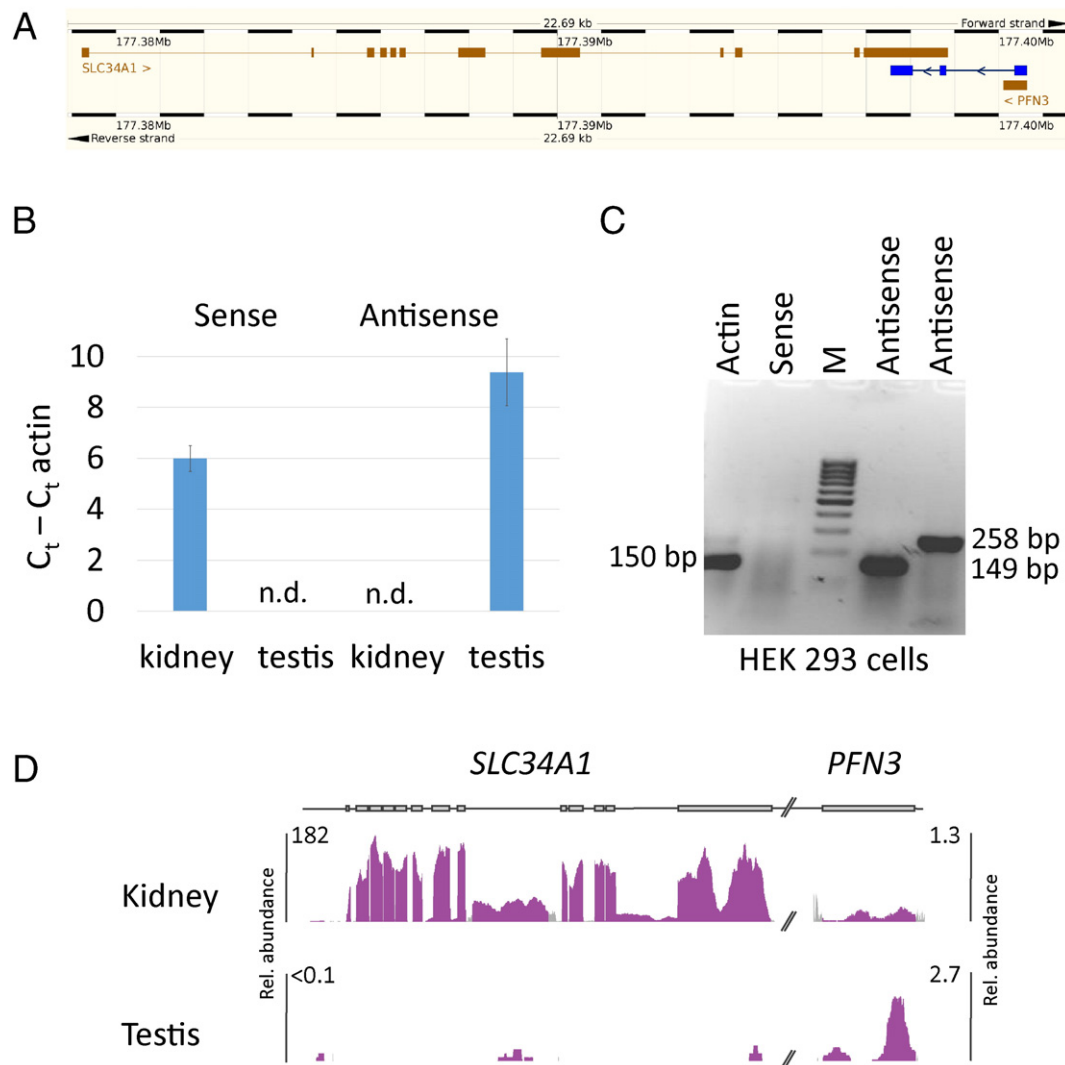


Fig. 1. The natural antisense transcript of *SLC34A1*. A) Genomic arrangement of the *SLC34A1* antisense transcript, representing an alternative splice form of *PFN3* (in blue). B) RT-qPCR of *SLC34A1* sense and antisense transcripts. C) RT-PCR with *SLC34A1* specific primers using HEK293 total RNA. Two different primer pairs for the antisense transcript were used (Supplementary Table 1). D) Representation of RNA-Seq data from human protein atlas.

in fish) have different origins: The *Slc34a1* related NATs derive from the *Pfn3* promoter and this gene is only found in mammals and marsupials. Antisense transcripts complementary to *Slc34a2* from zebrafish are thought to be driven by the bi-directional *Rbpja* promoter. The NAT-relevant promoter in winter flounder has not been defined [19]. The arrangement of genes around flounder *Slc34a2* suggests that the antisense transcript is either an alternatively spliced *Sel1l3* product or derives from an undefined promoter between *Slc34a2* and *Sel1l3*. The apparently rather random emergence of the *Slc34a* antisense transcripts in contrast to the well-conserved sense transcript and the established physiological role of the related transport proteins implies that the sense and antisense transcripts follow distinct evolutionary agendas. Specifically, phosphate is essential for growth and its levels are tightly regulated, in vertebrates by hormonally controlled expression of *Slc34a* encoded proteins. At any point, testis is neither directly nor indirectly involved in maintaining phosphate homeostasis. This finding raises two questions; first, if the antisense transcripts is unlikely to modulate the physiological functioning of the sense encoded protein, what else does it do; and second, to which extent is the situation in *Slc34a* relevant to other bi-directionally transcribed genes?

3.3. Conservation of bi-directionally transcribed loci

The expression pattern as well as mechanistic considerations can provide clues to identify putative biological functions of sense and antisense transcript pairs. There are a number of mechanisms relating to natural antisense transcripts and particularly double-stranded RNA which have been discussed extensively and reviewed in detail [23,24]. RNA editing, RNA interference and RNA masking involve dsRNA intermediates whereas transcriptional interference and chromatin changes may not require RNA-RNA pairing. We selected 10 bi-directionally transcribed loci that were previously characterized in detail and collected a number of coordinates from public databases including coding potential of the transcripts, expression patterns and phylogenetic conservation (Table 1). The selection is rather small because investigations into the mechanism of gene regulation by NATs in animals are limited. Nevertheless, the data confirm published findings that NATs are generally expressed at a low level poorly conserved. The observation that only one of 14 antisense transcripts is structurally conserved (intron/exon structure, not just antisense transcription per se) suggests that the overall structure of these transcripts is not essential for a potential biological function. All the reports that included mechanistic studies found

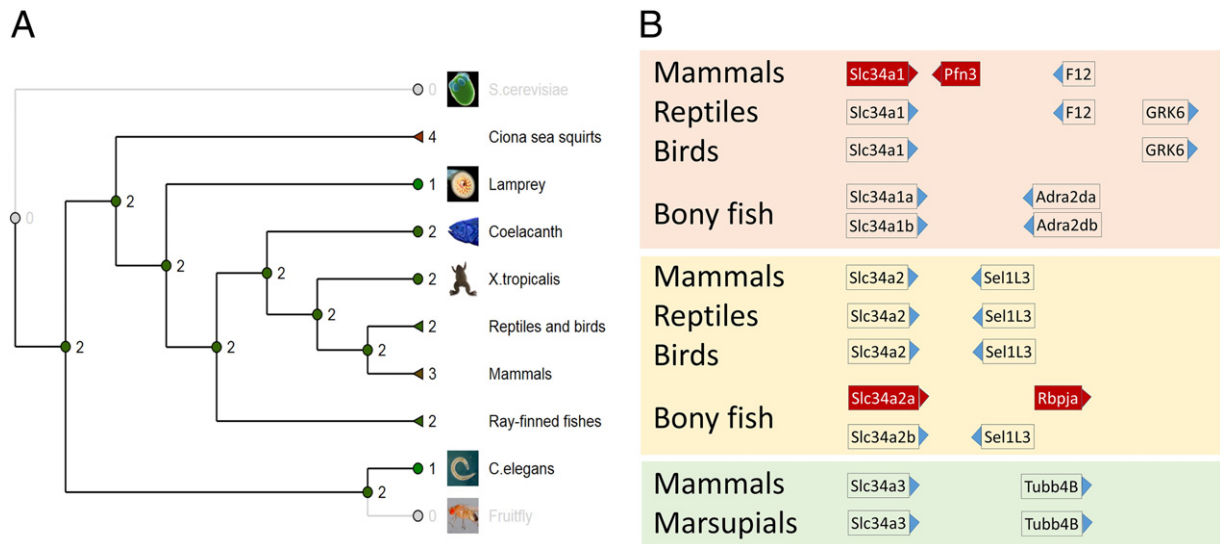


Fig. 2. Phylogenetic tree of *Slc34A* genes of Na-phosphate transporters. A) Phylogenetic tree with the number of isoforms indicated. B) The tree isoforms of mammalian *Slc34A*, *Slc34A2* (peach), *Slc34A3* (yellow) and *Slc34A3* (green), and the genes downstream that are relevant to potential antisense transcription. *Slc34A* isoforms including the downstream genes with confirmed expression of antisense transcripts are in red.

evidence of dsRNA formation between sense and antisense transcripts; those who did not, found either no exonic complementarity within the locus (*RUNX2/SUPT3H*) or did not investigate the question. Moreover, the expression profile of the various transcripts suggested an accumulation of lowly expressed NATs in testis. The importance of a dsRNA intermediates is corroborated by a recent report that detected significant levels of sense-antisense hybrids in human cells [12].

In order to test the importance of dsRNA formation at a medium scale we compiled a random list of 226 human bi-directionally transcribed genes. The only requirement was complementarity of 20 or more base pairs of the primary transcripts. These transcripts were then catalogued with respect to their genomic arrangement, exonic complementarity and expression pattern (Supplementary Table 2). As demonstrated in Fig. 3B, natural antisense transcripts were found to be predominantly non-protein-coding (72.6%, compare column 1 and column 3). The genomic arrangement showed approximately equal proportions of 3', mid and 5' overlapping antisense transcripts (Fig. 3C). The proportion of antisense transcripts with exonic complementarity (NATs) to purely intronic transcripts (NOTs) was about 2:1 (62%/38%). Complementarity between sense and antisense transcripts was predominantly found at the 5' and the 3' end of genes (first and last exon, respectively), whereas antisense transcripts embedded within the sense encoding gene tend to be non-overlapping (NOTs) (Fig. 3C, compare columns 1 and 2). Our findings are in line with previously published large scale assessments of the human antisense transcriptome [3].

We then assessed the phylogenetic conservation of the human loci in mouse and compiled the same set of parameters (genomic arrangement and expression pattern). Synteny was found for all but one locus (*MMP26*) and most of the (sense) transcripts show a similar expression levels in both human and mouse (Fig. 3A, $r = 0.851$ in Supplementary Fig. 1). In contrast, antisense transcript expression was neither correlated between human and mice nor in relation to the cognate sense transcripts (Supplementary Fig. 1). Only 50% (113) of the syntenic mouse loci showed evidence of bi-directional transcription (Fig. 3B, columns 3 and 4). As expected, protein coding antisense transcripts were more likely to be conserved (91.9% protein coding versus 34.1% non-protein coding, respectively, Fig. 3B, columns 3 and 4), particularly at the 5' and 3' end of genes: 58.5 and 63.7% versus 32.6% mid gene (Fig. 3C, columns 3 and 4). When the genomic organisation was taken into account, only 37 (33%) antisense transcripts displayed comparable features (exonic complementarity/location of the antisense transcript relative to the

sense transcript) between mouse and human. The summarized expression pattern in brain/testis/other tissues of the transcripts was comparable between mouse and human and, for both species antisense transcripts were enriched in testis (Fig. 4A).

The significant expression of antisense transcripts in testis points to a biological function of – at least some – antisense transcripts that is not directly related to the sense encoded protein function. Our results also confirm previous reports that NATs are structurally not well conserved between species [25].

3.4. Antisense transcripts on the X chromosome

In an evolutionary context, bi-directionally transcribed genes on human and murine X chromosomes are of particular interest: Transposable elements, which drive the evolution of new genetic elements and are also linked to the emergence of antisense transcripts [26] are significantly enriched on mammalian sex chromosomes [27]. Despite the high transposon density the mammalian X chromosome show a significant reduction in NATs as compared to the density on autosomes [3,4]. We argued that evolutionary conservation of a complementary gene arrangement on the X chromosome (where antisense transcripts are negatively selected) would provide support for a regulatory impact of the antisense transcript on the sense-encoded protein function. We catalogued bi-directionally transcribed loci on both human and mouse X chromosome as described above. We found 133 (16.1% of X-linked genes) bi-directionally transcribed loci in human and 93 (9.9% of X-linked genes) in mouse, respectively. These findings confirm the previously reported bias against complementary transcripts on mammalian X chromosomes. Interestingly, the expression pattern of sense transcripts was comparable to autosomal genes in both human and mice; X-encoded antisense transcripts, however, showed a tendency against expression in testis (contrary to autosomal encoded NATs, Fig. 4B). Synteny was detected for about 90% of the bi-directionally transcribed loci in both human-mouse (83.5%) and mouse-human direction (91.4%; Fig. 5A, compare the sum of the upper two panels with the bottom panel), which is lower than the almost complete (>99.5%) synteny of autosomal genes. Bi-allelic transcription was conserved in 42 loci (Fig. 5A, top) 68 loci are specific to human, 43 to mouse (Fig. 5A, middle panel). In human, the proportion of non-protein coding to coding antisense transcripts was 35.8%, in mouse 51.4% respectively (sum of grey bar areas in Fig. 5A). Analysis of the genomic arrangement revealed a clear bias of complementarity between sense and antisense at the 3'

Table 1
Compilation of well-studied bi-directionally transcribed genes. Peach-coloured transcripts are protein coding, yellow indicates a mode of action that involves dsRNA formation. Expression levels and patterns are deduced from the EBI expression atlas (<http://www.ebi.ac.uk/gxa/home>). The proposed mechanism is informed by the indicated papers.

Gene name	Expression	Antisense gene name	Expression	Overlap	Structure	Expression AS/S	Expression pattern sense	Expression pattern antisense	Proposed mechanism	Synteny human-mouse, sense	Synteny human-mouse, antisense	
BACE1 (8)	47	BACE1-AS	12	Mid	Multiple exons	NAT	0.26	Brain, pancreas	Brain, multiple	miRNA masking	Conserved	nd
BACE1	47	RNF214	14	5'	1st exon	NAT	0.30	Brain, pancreas	Testis, multiple	Not determined	Conserved	Not conserved
WDR83	16	WDR83-OS	259	Mid	Multiple exons	NAT	16.19	Brain, testis, multiple	Multiple	Not determined	Conserved	Not conserved
WDR83 (27)	16	DHPS	118	5'	1st exon	NAT	7.38	Brain, testis, multiple	Brain, multiple	dsRNA + stability	Conserved	Not conserved
DHRS4 (17)	49	DHRS4-AS1	22	Mid, 5'	1st exon	NAT	0.45	Multiple	Multiple	dsRNA, Me/histone	nd	nd
HAS2 (20, 30)	12	HAS2-AS	3	5'	1st exon	NAT	0.25	Digestive tract, testis	Testis, digestive tract	dsRNA, + stability/ chromatin	Conserved	nd
HIF1a (29)	184	HIF1a-AS1	4	5'	1st exon	NAT	0.02	Multiple	Multiple	+ stability	Conserved	nd
TP53 (23)	63	WRAP53	5	5'	1st exon	NAT	0.08	Multiple	Testis, multiple	dsRNA, + stability	Conserved	Not conserved
PINK1 (24, 25)	31	PINK1-AS	3	3'	Multiple exons	NAT	0.10	Testis, multiple	Multiple	Concordant regulation	Conserved (?)	nd
RUNX2 (1)	7	SUPT3H	62	5'	Intronic	NOT	8.86	Immune system, epithelia	Brain, multiple	Promoter interaction	Conserved	conserved
RUNX2	7	RP1-244F24	3	Mid	Intronic	NOT	0.43	Immune system, epithelia	Multiple, glands	Not determined	Conserved	nd
SCN1A	14	AC010127.3	1	Mid	Multiple exons	NAT	0.07	Brain	Testis		Conserved	nd
SCN9A (16)	4	AC010127.3	1	Mid	Multiple exons	NAT	0.25	Brain	Testis	Discordant	Conserved	nd
SLC34A1 (5)	219	PFN3	39	3'	Multiple exons	NAT	0.19	Kidney	Testis	siRNA, discordant	Conserved	Not conserved\

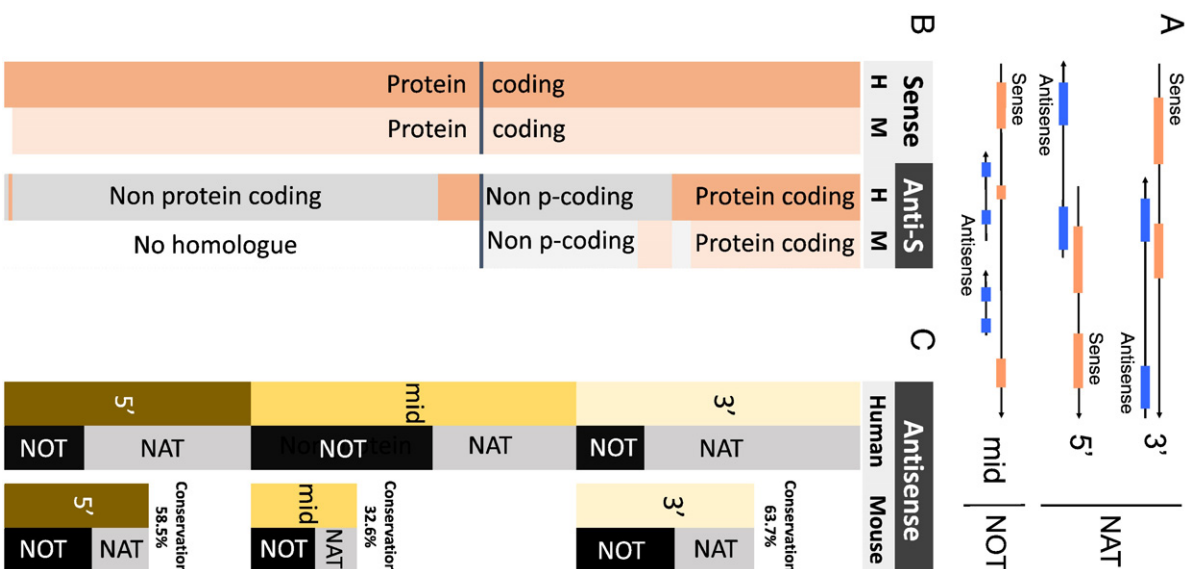


Fig. 3. Natural antisense transcripts on autosomes of human (H) and mouse (M). A) Classification of sense-antisense overlapping structures including 3' (tail-to-tail), mid (full overlap) and 5' (head-to-head). Inclusion into the dataset required complementarity of >20 nucleotides. B) Expression pattern of sense and antisense transcripts and their phylogenetic conservation. Peach and light peach indicate protein coding potential in human and mouse, respectively. Protein coding sense transcripts are well conserved between human and mouse (left two columns), the antisense transcripts to a much lesser extent (right columns) B) Comparison of the genomic arrangement of bi-directionally transcribed genes in human and mouse. 3', mid and 5' indicates the position of complementarity between sense and antisense transcripts. NAT stands for an antisense transcript with complementarity to the sense RNA, NOT indicates a non-overlapping configuration. S = sense, AS = antisense.

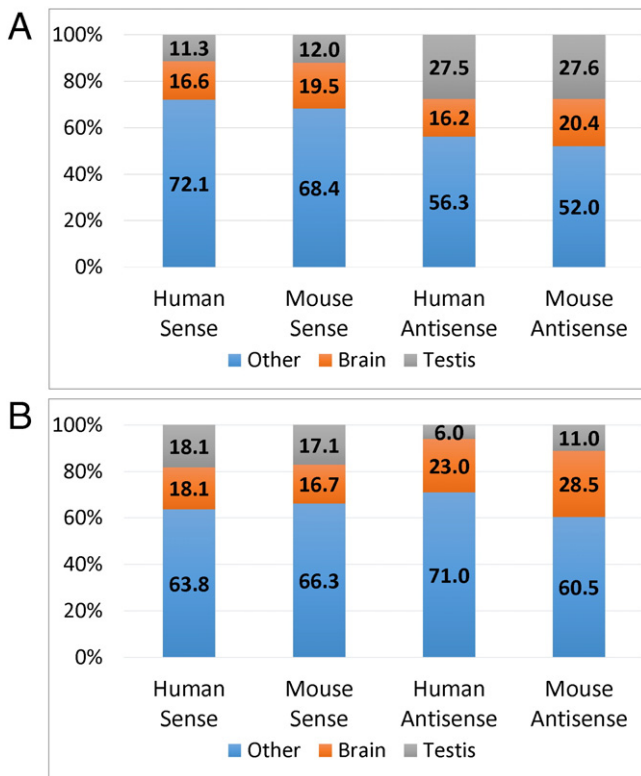


Fig. 4. Schematic expression pattern of sense and antisense transcripts in human and mice. Expression data of individual genes was extracted from the gene expression atlas (<http://www.ebi.ac.uk/gxa/home>), scored and grouped into brain, testis and other tissues. A) Genes on autosomes including sense and antisense transcripts, B) Bi-directional X-chromosomal genes of human and mice.

chromosome inactivation (37, 3.9%) and only 4 were found to be bi-directional and escapee, too few to warrant conclusions.

In summary, our findings point to selective benefits to retain antisense transcripts but also frequent novel incidences of bi-directional transcription in human and mouse lineages [29]. Moreover, the reduced number of bi-directionally transcribed genes on X chromosomes show lesser conservation than autosomal loci (31.6% vs. 50%).

4. Discussion

The scope of natural antisense transcripts is still largely speculative and often indirect evidence such as phylogenetic conservation is used to underpin biological relevance. The task of investigating functionality of NATs is also hampered by the fact that experimental up- or down regulation of either transcript of a sense-antisense pair will almost inevitably interfere with the expression of the counterpart. Therefore, such experiments need to be very carefully controlled and interpreted.

The cloning of a human antisense transcript of the *SLC34A1/PFN3* locus that we report in this paper highlights another difficulty in the 'antisense' field: Despite ever increasing sequencing depth, the annotation of the antisense transcriptome is still incomplete. The generally low expression of antisense transcripts certainly exacerbates the problem. The material used for RNA extraction is another important factor to be considered; antisense transcripts are predominantly expressed in testis and systematic assessment of bi-directional genes should involve this tissue [30]. Moreover, unless the experimental strategy is aimed at maintaining strand specificity the bioinformatic assessment of RNA-seq data will have a bias against antisense transcripts because antisense reads are indiscernible from reverse transcription artefacts and DNA contamination.

We report the cloning of an alternative splice product of *PFN3* that is complementary to the last exon of *SLC34A1*. Since the two transcripts show a non-overlapping expression pattern, the sense transcript is almost exclusively expressed in kidney whereas the antisense transcript is found in testis, it is unlikely that the latter contributes to the physiological regulation of the Na-phosphate transporter and hence helps maintaining acute phosphate balance. It could, however, be involved in setting chromatin marks in kidney but also during sperm development. We have recently proposed a mechanism based on double stranded RNA formation and *endo*-siRNA processing in a perspective paper that may explain how epigenetic modifications can be targeted either to the sense or the antisense strand [31]. In that respect it is intriguing, that the kidney cell line HEK293 expresses the antisense transcript but lacks the message for the transporter.

SLC34A-encoded transporters play a central role in maintaining phosphate homeostasis in most animals as demonstrated in various animal models and knock-out experiments [32]. Accordingly, *SLC34A* isoforms from various species have been functionally assessed and shown to have adapted to specific physiologic/environmental constraints [33]. Bi-directional transcription has been detected in some isoforms but not in others; moreover, there is heterogeneity with respect to the phylogeny of *SLC34A*-related antisense transcripts: In fish they are driven by a bi-directional (*Rbpja*) promoter and in mammals they represent a read-through transcript of *Pfn3*. The unrelated occurrence of NATs in two different *SLC34A* isoforms suggests that the transcripts are unlikely to contribute to the acute regulation of Na-phosphate transporters. In general, the tissue expression of homologous (protein-coding) sense transcripts is well conserved between human and mouse but not so much the expression pattern of the cognate antisense transcripts (Fig. 3A, Supplementary Fig. 1, Supplementary Table 2). This observation provides additional support to the hypothesis that many antisense transcripts are not involved in the acute physiological regulation of the sense transcript.

The phylogenetic conservation of bi-directional loci between human and mouse has been investigated by others with comparable results [34, 35]. The figures crucially depend on the specific parameters to identify and categorize natural antisense transcripts but also on sequencing depth and annotation. The current deficit of mechanistic insights further complicates a comparative evaluation of antisense transcription. We focused on transcripts with at least 20 nucleotides of complementarity and only included protein coding and non-protein coding processed transcripts but excluded pseudogenes. Bi-directional transcription was conserved between human and mouse in about half of all investigated genes, in good agreement with other reports. Likewise, 3' and 5' complementary transcripts were better conserved than intronic, non-overlapping transcripts (NOTs). The slightly enhanced phylogenetic stability of complementary sense-antisense arrangements suggest that double-stranded RNA formation may represent an essential intermediate in an antisense mediated regulatory mechanism, a hypothesis that receives support from recently published mechanistic studies (Table 1).

The under-representation of antisense transcripts and overlapping gene arrangements becomes immediately apparent upon visual inspection of an annotated X chromosome in a genome browser [3,4]. Comprehensive analysis revealed a slightly higher percentage of bi-directional loci on the human than the mouse X chromosome. The conservation of X-encoded antisense loci was slightly lower as compared to genes on autosomes (human → mouse 32.6%, mouse → human 45.7% versus 50% human → mouse on autosomes), this may reflect a slightly accelerated evolution of the X chromosome as compared to autosomes [36].

X chromosome inactivation in human and mouse show significant mechanistic differences that also result in a significantly higher number of human genes that escape silencing as compared to mouse [37]. Interestingly, we found a significant overlap between X inactivation escape and complementary antisense transcripts in human (potentially forming sense-antisense RNA hybrids). In mouse, numbers were too

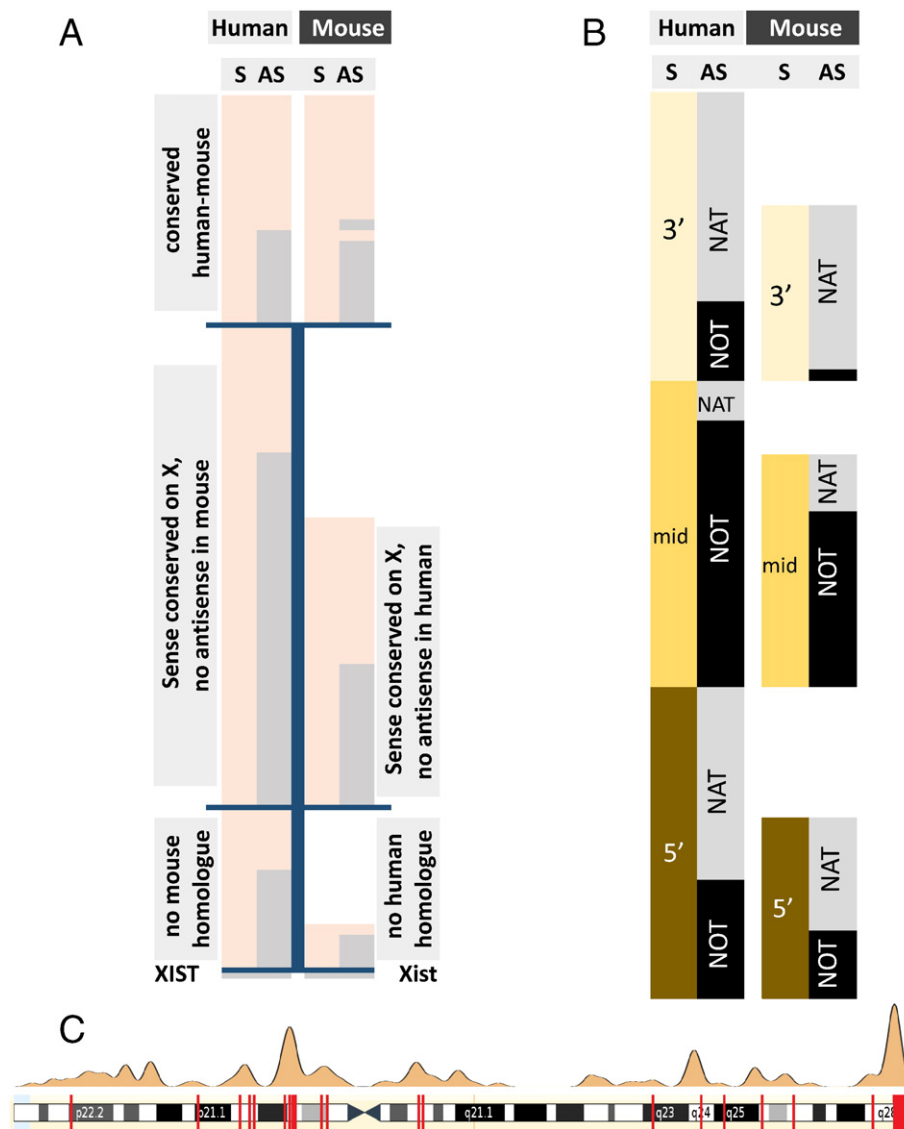


Fig. 5. Natural antisense transcripts on X chromosomes of human and mouse. A) Phylogenetic conservation of sense and antisense transcripts. Top panel, conserved bi-directional locus between human and mouse; middle panel, locus conserved but no antisense transcript in the other species; homologue in the other species. Light peach indicates protein coding potential, S = sense, AS = antisense. The nature of synteny is described in the grey boxes. B) Comparison of the genomic arrangement of bi-directionally transcribed genes on the X chromosome of human and mouse. 3', mid and 5' indicates the position of complementarity between sense and antisense transcript. NAT stands for an antisense transcript with complementarity to the sense RNA; NOT indicates a non-overlapping configuration. C) Natural antisense transcripts in regions that escape X chromosome inactivation. Upper panel indicates areas with relaxed inactivation, the red bars on the lower panel depict the bi-directionally transcribed human genes.

low to draw significant conclusions. These findings suggest that anti-sense transcription carries the potential of inducing transcriptional silencing which is potentially detrimental for already monoallelically expressed (X chromosomal) gene [38].

5. Conclusions

We have identified a novel antisense transcript complementary to the human *SLC34A1* gene. Expression pattern of the sense and antisense transcripts argue against an involvement of the antisense RNA in the acute regulation of the sense encoded Na-phosphate transporter. Comparative studies between human and mouse bi-directionally transcribed loci indicate generally limited conservation of antisense transcripts. Moreover, one of potential mechanisms involves dsRNA formation that may lead to the establishment of an epigenetic signature and potentially transcriptional gene silencing [38,39].

Comparative studies between human and mouse have revealed distinct expression signatures of (autosomal and X chromosomal) bi-

directionally transcribed loci in testis, indicating that the expression pattern established for *Slc34a1* transcripts is not unique. The particular importance of testis in the context of antisense-mediated gene regulation is also supported by expression differences of RNA interference enzymes in testis and somatic tissues [40,41]. Our results did not identify a bias in human-mouse conservation of overlapping versus non-overlapping sense-antisense transcripts; hence the comparative findings do not strengthen the argument for a mechanism involving dsRNA. The various hallmarks in genomics and expression of antisense transcripts, however, suggest tissue specific biological roles that do not necessarily relate to the sense encoded protein function.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ygeno.2016.05.004>.

Acknowledgment

This work was funded by The Dunhill Medical Trust (SA10/0210).

References

- [1] T. Beiter, E. Reich, R.W. Williams, P. Simon, Antisense transcription: a critical look in both directions, *Cell. Mol. Life Sci.* 66 (2008) 94–112.
- [2] S. Katayama, Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, C.C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C. Wells, M. Frith, T. Ravasi, K.C. Pang, J. Hallinan, J. Mattick, D.A. Hume, L. Lipovich, S. Batalov, P.G. Engstrom, Y. Mizuno, M.A. Faghihi, A. Sandelin, A.M. Chalk, S. Mottagui-Tabar, Z. Liang, B. Lenhard, C. Wahlestedt, Antisense transcription in the mammalian transcriptome, *Science* 309 (2005) 1564–1566 (New York, N.Y.).
- [3] J. Chen, M. Sun, W.J. Kent, X. Huang, H. Xie, W. Wang, G. Zhou, R.Z. Shi, J.D. Rowley, Over 20% of human transcripts might form sense-antisense pairs, *Nucleic Acids Res.* 32 (2004) 4812–4820.
- [4] H. Kiyosawa, I. Yamanaka, N. Osato, S. Kondo, Y. Hayashizaki, Antisense transcripts with FANTOM2 clone set and their implications for gene regulation, *Genome Res.* 13 (2003) 1324–1334.
- [5] R. Yelin, D. Dahary, R. Sorek, E.Y. Levanon, O. Goldstein, A. Shoshan, A. Diber, S. Biton, Y. Tamir, R. Khosravi, S. Nemzer, E. Pinner, S. Walach, J. Bernstein, K. Savitsky, G. Rotman, Widespread occurrence of antisense transcription in the human genome, *Nat. Biotechnol.* 21 (2003) 379–386.
- [6] E.G. Wagner, S. Brantl, Kissing and RNA stability in antisense control of plasmid replication, *Trends Biochem. Sci.* 23 (1998) 451–454.
- [7] E.G. Wagner, P. Romby, Small RNAs in bacteria and archaea: who they are, what they do, and how they do it, *Adv. Genet.* 90 (2015) 133–208.
- [8] Q. Wang, G.G. Carmichael, Effects of length and location on the cellular response to double-stranded RNA, *Microbiol. Mol. Biol. Rev.* 68 (2004) 432–452.
- [9] S.W. Ding, O. Voinnet, Antiviral immunity directed by small RNAs, *Cell* 130 (2007) 413–426.
- [10] M.A. Garcia, E.F. Meurs, M. Esteban, The dsRNA protein kinase PKR: virus and cell control, *Biochimie* 89 (2007) 799–811.
- [11] L.F. Luo, C.C. Hou, W.X. Yang, Small non-coding RNAs and their associated proteins in spermatogenesis, *Gene* 578 (2016) 141–157.
- [12] M.M. Portal, V. Pavet, C. Erb, H. Gronemeyer, Human cells contain natural double-stranded RNAs with potential regulatory functions, 22 (2015) 89–97.
- [13] A. Werner, S. Cockell, J. Falconer, M. Carlile, S. Alnumeir, J. Robinson, Contribution of natural antisense transcription to an endogenous siRNA signature in human cells, *BMC Genomics* 15 (2014) 19.
- [14] K.H. Ling, P.J. Brautigan, S. Moore, R. Fraser, P.S. Cheah, J.M. Raison, M. Babic, Y.K. Lee, T. Daish, D.M. Mattiske, J.R. Mann, D.L. Adelson, P.Q. Thomas, C.N. Hahn, H.S. Scott, Derivation of an endogenous small RNA from double-stranded Sox4 sense and natural antisense transcripts in the mouse brain, *Genomics* 107 (2016) 88–99.
- [15] X. Ge, W.S. Rubinstein, Y.C. Jung, Q. Wu, Genome-wide analysis of antisense transcription with Affymetrix exon array, *BMC Genomics* 9 (2008) 27.
- [16] A. Werner, G. Schmutzler, M. Carlile, C.G. Miles, H. Peters, Expression profiling of antisense transcripts on DNA arrays, *Physiol. Genomics* 28 (2007) 294–300.
- [17] A. Werner, R.K. Kinne, Evolution of the Na-P(i) cotransport systems, *Am. J. Phys. Regul. Integr. Comp. Phys.* 280 (2001) R301–R312.
- [18] M. Carlile, D. Swan, K. Jackson, K. Preston-Fayers, B. Ballester, P. Flicek, A. Werner, Strand selective generation of endo-siRNAs from the Na/phosphate transporter gene Slc34a1 in murine tissues, *Nucleic Acids Res.* 37 (2009) 2274–2282.
- [19] B. Huelsseweh, B. Kohl, H. Hentschel, R.K. Kinne, A. Werner, Translated anti-sense product of the Na/phosphate co-transporter (NaPi-II), *Biochem. J.* 332 (Pt 2) (1998) 483–489.
- [20] P. Nalbant, C. Boehmer, L. Dehmelt, F. Wehner, A. Werner, Functional characterization of a Na + -phosphate cotransporter (NaPi-II) from zebrafish and identification of related transcripts, *J. Physiol.* 520 (Pt 1) (1999) 79–89.
- [21] Y. Zhang, A. Castillo-Morales, M. Jiang, Y. Zhu, L. Hu, A.O. Urrutia, X. Kong, L.D. Hurst, Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving, *Mol. Biol. Evol.* 33 (2016) 302.
- [22] J.B. Berletch, W. Ma, F. Yang, J. Shendure, W.S. Noble, C.M. Disteche, X. Deng, Escape from X inactivation varies in mouse tissues, *PLoS Genet.* 11 (2015), e1005079.
- [23] M. Magistri, M.A. Faghihi, G. St Laurent 3rd, C. Wahlestedt, Regulation of Chromatin Structure by Long Noncoding RNAs: Focus on Natural Antisense Transcripts (Trends Genet) 2012.
- [24] A. Werner, M. Carlile, D. Swan, What do natural antisense transcripts regulate? *RNA Biol.* 6 (2009) 43–48.
- [25] V. Veeramachaneni, W. Makalowski, M. Galdzicki, R. Sood, I. Makalowska, Mammalian overlapping genes: the comparative perspective, *Genome Res.* 14 (2004) 280–286.
- [26] S. Ayarpadikannan, H.E. Lee, K. Han, H.S. Kim, Transposable element-driven transcript diversification and its relevance to genetic disorders, *Gene* 558 (2015) 187–194.
- [27] F.F. Costa, Non-coding RNAs, epigenetics and complexity, *Gene* 410 (2008) 9–17.
- [28] H. Kiyosawa, K. Abe, Speculations on the role of natural antisense transcripts in mammalian X chromosome evolution, *Cytogenet. Genome Res.* 99 (2002) 151–156.
- [29] E.M. Kvikstad, K.D. Makova, The (r)evolution of SINE versus LINE distributions in primate genomes: sex chromosomes are important, *Genome Res.* 20 (2010) 600–613.
- [30] M. Soumillon, A. Necseulea, M. Weier, D. Brawand, X. Zhang, H. Gu, P. Barthes, M. Kokkinaki, S. Nef, A. Gnirke, M. Dym, B. de Massy, T.S. Mikkelsen, H. Kaessmann, Cellular source and mechanisms of high transcriptome complexity in the mammalian testis, *Cell Rep.* 3 (2013) 2179–2190.
- [31] A. Werner, M.J. Piatek, J.S. Mattick, Transpositional shuffling and quality control in male germ cells to enhance evolution of complex organisms, *Ann. N. Y. Acad. Sci.* 1341 (2015) 156–163.
- [32] J. Biber, N. Hernandez, I. Forster, Phosphate transporters and their function, *Annu. Rev. Physiol.* 75 (2013) 535–550.
- [33] I.C. Forster, N. Hernandez, J. Biber, H. Murer, Phosphate transporters of the SLC20 and SLC34 families, *Mol. Asp. Med.* 34 (2013) 386–395.
- [34] E.J. Wood, K. Chin-Inmanu, H. Jia, L. Lipovich, Sense-antisense gene pairs: sequence, transcription, and structure are not conserved between human and mouse, *Front. Genet.* 4 (2013) 183.
- [35] P.A. Galante, D.O. Vidal, J.E. de Souza, A.A. Camargo, S.J. de Souza, Sense-antisense pairs in mammals: functional and evolutionary considerations, *Genome Biol.* 8 (2007) R40.
- [36] J.M. Turner, Meiotic silencing in mammals, *Annu. Rev. Genet.* 49 (2015) 395–412.
- [37] A.V. Gendrel, E. Heard, Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation, *Annu. Rev. Cell Dev. Biol.* 30 (2014) 561–580.
- [38] C. Tufarelli, J.A. Stanley, D. Garrick, J.A. Sharpe, H. Ayyub, W.G. Wood, D.R. Higgs, Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease, *Nat. Genet.* 34 (2003) 157–165.
- [39] W. Yu, D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A.P. Feinberg, H. Cui, Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA, *Nature* 451 (2008) 202–206.
- [40] E. Gonzalez-Gonzalez, P.P. Lopez-Casas, J. del Mazo, The expression patterns of genes involved in the RNAi pathways are tissue-dependent and differ in the germ and somatic cells of mouse testis, *Biochim. Biophys. Acta* 1779 (2008) 306–311.
- [41] A.J. Modzelewski, R.J. Holmes, S. Hilz, A. Grimson, P.E. Cohen, AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline, *Dev. Cell* 23 (2012) 251–264.